DUKE ENVIRONMENTAL AND ENERGY ECONOMICS WORKING PAPER SERIES
organized by the
NICHOLAS INSTITUTE FOR ENVIRONMENTAL POLICY SOLUTIONS
and the
DUKE UNIVERSITY ENERGY INITIATIVE

Comparing the Statistical Distributions of Energy Efficiency in Manufacturing: Meta–Analysis of 24 Industry-Specific Case Studies

Gale Boyd*
Working Paper EE 15-03
July 2015

* Social Science Research Institute, Duke University, Box 90989, Durham, NC 27708; gale.boyd@duke.edu

Acknowledgments

This paper was developed at Duke University with funding from the U.S. Environmental Protection Agency's Office of Atmospheric Programs, ENERGY STAR for Industry. The paper and associated analysis would not have been possible without the input of industry participants in the ENERGY STAR Focuses on Energy Efficiency. Their willingness to provide data, guidance on issues affecting manufacturing energy use, and reviews of case studies was invaluable. Thanks also go to Walt Tunnessen, Betsy Dutrow, Josh Smith, and the reviewers and participants of the ACEEE Summer Study on Energy Efficiency in Industry, July 2013, for their comments on an earlier version of this paper. Some results in this paper were prepared while the author was a special sworn status research associate at the Triangle Census Research Data Center. Any opinions and conclusions expressed herein are those of the author. All results have been reviewed to ensure that no confidential information is disclosed. Any errors are the sole responsibility of the author. Many research assistants were involved in the preparation of these case studies. Only some are cited as co-authors in the reference section; others include Tatyana Kuzmenko, Béla Személy, Chris Geissler, Jeremy Chiu, Songman Kang, and Shouyue Yu.

The Duke Environmental and Energy Economics Working Paper Series provides a forum for Duke faculty working in environmental, resource, and energy economics to disseminate their research. These working papers have not necessarily undergone peer review at the time of posting.

ABSTRACT

Over the past several years, there has been growing interest among policy makers and others in the role that industrial energy efficiency can play in climate, air, and other potential regulatory actives. For over ten years, the U.S. Environmental Protection Agency (EPA) has supported the development of sector specific industrial energy efficiency case studies using statistical analysis of plant level data to assess the distribution of energy use, controlling for a variety of plant production characteristics. These case studies are the basis for the ENERGY STAR® Energy Performance Indicators (EPI). To date there are EPI for fourteen broad industries, two dozen sectors, and many more detailed product types. This paper is a meta-analysis of the approach that has been used in this research and the general findings regarding the range of performance within and across industries. Observations about industrial plant benchmarking and lessons learned are explored. We find that there are few sectors that are well represented by a simple "energy per widget" benchmark; that less energy intensive sectors tend to exhibit a wider range of within industry performance than energy intensive sectors; and that changes over time in the level and range of energy performance do not reveal any single pattern.

Introduction

ENERGY STAR is a voluntary program launched by the EPA in 1992 to identify and promote energy efficient products, buildings, homes, and manufacturing facilities.[1] The program was established to find cost-effective ways to reduce greenhouse gas emissions associated with energy use. Initially focused on consumer products, the program expanded into the commercial building market in 1995 and released its first energy-efficiency benchmark for office buildings in 1998. In 2000, the EPA expanded the program to include manufacturing plants. ENERGY STAR focuses on providing tools that encourage better corporate energy management through the development of sector specific energy performance benchmarks. One goal is to provide companies within a manufacturing sector with an objective measure of how their plant compares to the rest of the industry. Most companies lack sufficient information on the relative efficiency of their plants within the broader industry because that information is confidential. Consequently, many companies did not know if their companies where operating efficiently or where the frontier for improved efficiency lies. The second objective of these studies was to establish criteria for determining which plants would qualify for recognition. In order to provide recognition for energy efficient manufacturing plants, EPA needed an objective and transparent means to determine which plants are best-in-class, defined by ENERGY STAR as *"the upper quartile of performance for similar production facilities."*

To develop this, one needs to address a variety of key issues:

- How to define energy efficiency?
- How do you control for differences between plants?
- What statistical distributions can be used to measure the quartiles?

In addition to these basic questions, data is key to the development of these case studies. This paper is an overview and meta-analysis of the case study approach used based on non-public, plant-level data from the U.S. Census Bureau, Triangle Research Data Center and other non-public sources.

Defining Energy Efficiency[2]

Efficiency is a measure of relative performance; but relative to what? Defining energy efficiency requires a choice of a reference point or benchmark against which to compare energy use. Energy efficiency benchmarks can be developed through a variety of means, such as engineering and theoretical estimates of performance or through observing the range of actual levels of performance. The choice of method used to define efficiency depends on the need to define a reference point for energy efficiency. One of the challenges with using energy efficiency benchmarks based on engineering or theoretical estimates is that they are often dismissed by industry as being economically infeasible. Consequently, these case studies has focused on developing benchmarks based on actual or observed operational performance rather

[1] See EPA (2011) for more information.

[2] This section draws from Boyd, G. (2012). A Statistical Approach to Plant-Level Energy Benchmarks and Baselines: The Energy Star Manufacturing-Plant Energy Performance Indicator. Carbon Management Technology Conference. Orlando, Florida USA.

than theoretical estimates of potential efficiency levels. Additionally, EPA needed to identify a benchmarking method that would be perceived by users as providing economically feasible performance targets.

The reference point for economic potential (observed practice) depends, in part, on the reason for measuring efficiency as well as the available information to create a reference. Generally, the *Ceteris Paribus* principle ("all other things being equal or held constant") is usually desired in creating the reference point, or benchmark. From a practical perspective there is a hierarchy of measures and methods by which one can "hold constant" things that influence *energy use* that are not part of *energy efficiency.* The first is some measure of production activity, either production of the final product or, alternatively, a ubiquitous input into the production process. This is most commonly done by computing the ratio of energy use to activity, a measure of energy intensity. Energy intensity is a common metric that controls for changes in production and is commonly confused with energy efficiency, as in the statement "*the industry or plant's energy efficiency has improved based on the fact that the corresponding energy intensity has declined over time.*" This type of statement brings us to the second way that one may approach the ceteris paribus principle for measuring efficiency, comparing energy intensity a particular plant, firm, or industry to itself over time. This approach is a plant[3] specific *baseline* comparison, or *intra-plant* efficiency benchmark. The baseline approach has the advantage of controlling for some plant specific conditions that do not change during the comparison period.

The next level of this ceteris paribus principle is an *inter-plant* comparison that may include a variety of factors that influence energy use, but may not be viewed as efficiency. Factors may include difference in the types of product and materials used, as well as location specific conditions. Intra-plant comparisons within an industry also get us closer to the notion of an observed best-practice benchmark of economic energy efficiency, since by definition there is some group of plants that are the best performers. This was the notion introduced by (Farrell 1957) and has been the basis for measuring production efficiency in economics. A modified approach has been adopted (Boyd 2005) and its evolution is discussed by (Boyd, Dutrow et al. 2008).

Intensity Metric Selection

Intensity ratios provide a basic metric for measuring energy efficiency and performance compared to a baseline. To measure intensity you need a measure of energy and something for the denominator. For the numerator these studies use total source energy, defined as the net Btu total of the fuels (Btu) and electricity (Kwh) with electricity converted to Btu based on the level of efficiency of the U.S. grid for delivered energy, i.e. including generation and transmission losses. A net measure is needed for when energy is transferred off site, most commonly in the form of steam or electricity.

The choice of the denominator is a major issue for measuring intensity. Ideally the denominator should capture some measure production. (Freeman, Niefer et al. 1997) show that industry level trends in energy intensity based on value, both total and value added, can differ dramatically from those based on physical quantities. As Freeman et all have observed, there are

[3] Throughout the paper we will refer to the plant level as the unit of observation, but the concept may also apply to more aggregate levels like firms and industries, an sometimes to less aggregate levels like process units.

many challenges with creating efficiency benchmarks based on price indexes, cost and other value measurements.[4]

Given issues with linking energy use with price indexes, these studies have focused on using metrics based on physical quantities. For physical production to be meaningful it needs to be at a high level of industry specificity. For example, the "Dairy" industry produces many products that cannot be aggregated, but "Fluid Milk" can. Therefore, within industries, it is necessary to differentiate between specific types for plants and manufacturing operations.

Similarly, building energy use benchmarks commonly use physical size (ft^2) as the main denominator for energy intensity benchmarks, but for most industrial facilities this isn't appropriate.[5] While commonly used for commercial buildings where energy use is primarily tied to plug loads, lighting and HVAC systems, energy intensity based on size (sq. ft) does not correspond well with manufacturing process energy uses. While energy intensity ratios are commonly used for intra-plant level baseline comparisons in an industrial energy management setting, their value for developing inter-plant comparisons is limited. For inter-plant comparison, there are multiple factors that must be considered.

Multi-Factor Benchmarks

When making intra-plant comparisons, it necessary consider a variety of factors that do not neatly fit neatly under the denominator of an energy intensity ratio. While all plants may make a common product, other differences can significantly affect energy intensity. The difficulty with applying an industry level inter-plant benchmark is controlling for inter-plant difference other than production volume. While the things that differ between plants are numerous, we have found a common thread across industries that the primary difference that have the most impact on energy fall into the following categories.

- Product mix
- Process input choices (i.e. "make or buy" upstream integration)
- Size - Physical or productive capacity and utilization rates
- Climate (and other location specific factors)

The choice of factors to include in the analysis depends upon the nature of the production process, the configuration of the industry (e.g. is upstream integration common or rare), the availability of data to represent these factor, and ultimately the outcome of the statistical tests for significance. In order to address these types of factors, these studies use a multivariate approach to normalization where multiple effects are simultaneously considered (Boyd and Tunnessen 2007). The next sections discussed the four basic categories of effects that are commonly

[4] As Freeman et all (2007) note, *"For an industry producing a single, well-defined, homogeneous good, it is relatively easy to construct an accurate price index. Most industries, however, produce many poorly-defined, heterogeneous goods. For a variety of reasons, the more diverse the slate of products produced by an industry, the more difficult it becomes to construct an accurate price index. …the accuracy of industrial price indexes is of extreme importance to industrial energy analysts and policy makers who use value-based indicators of energy intensity."*

[5] The one exception is Pharmaceutical manufacturing where energy intensity is expressed as MMBTU/SQ FT. This metric was chosen largely because of the huge impact of HVAC systems in pharmaceutical manufacturing.

considered. There is further elaboration on the way this is implemented the section on industry specific comparisons.

Product Mix

Not all plants produce exactly the same product. In fact, many plants themselves produce multiple products. The diversity between plants gives rise to a mix of derived demands for specific processes and energy services. To the extent that the final product is the results of a series of energy using steps the energy use of the plant will depend on the level and mix of products produced. Rather than specifying each process step individually, the approach used here is to identify those products that use significantly more (or less) energy and measure those energy requirements with a statistical comparison.

One approach to controlling for product mix is to segment the industry into cohorts based on product categories. This works best when there is no overlap between plants that produce the various basic products and there are sufficient numbers of plants to conduct the statistical comparison between those resulting groups. This means each sub-group is effectively treated as a separate industry for evaluation proposes. A good example is the glass industry where containers and flat glass are distinct industry segments.

When such natural sub sectors do not exist and multiple products are produced within a plant, additional approaches are needed. The statistical approach is well suited to testing if a particular grouping of products is appropriate for benchmarking differences in energy. When industries produce a mix of products that differs across plants then the product mix (share of activity) of distinct products is needed. This approach was first used in wet corn mills (Boyd 2008) and was later applied to other sectors.

In the absence of meaningful data on discrete product classes an alternative is a continuous measure of product differentiation. Price is often taken as a measure of quality difference. To the extent that such quality difference arise for additional energy using processes then value of shipments may be an appropriate proxy for product mix. Differences in value may not involve higher energy use, as in luxury cars or specialty beers, but may be the case in creating different types of glass bottles or more complex cast metal products. Given available data the link between energy and value (price) can be treated as largely an empirical issue, but preferably with some underlying hypothesis about the industry in the case study. Value of shipments might be used instead of a physical production variable or in conjunction with physical outputs. In the latter case the ratio of value to physical product is price and becomes an implicit variable in the analysis. Other measures of energy related product differences are industry specific; as in the case of vehicle size in automobile assembly,

Size

Size and associated capacity utilization rates may directly impact energy use. Size may impact specific engineering and managerial advantages to energy use. If there is a substantial "fixed" level of energy use in the short run, the utilization rates may have a non-linear impact on energy intensity. In order to include size (and utilization) as a normalizing factor a meaningful measure of size or capacity is needed. It may be measured on an input basis, output basis, or physical size. In some cases there may be advantages to larger scale of production, i.e. economies of scale. If it is the case that a larger production capacity or larger physical plant size has less than proportionate requirements for energy consumption then there are economies of scale with respect to energy use. For example, in the cement industry the scale is quite important. The larger size of the kiln (rather than several smaller kilns) has advantages in terms of energy use. The analysis for this particular sector accounts for this.

Process Inputs

There are three ways that process inputs are important for benchmarking. The first is that inputs like materials, labor, or production hours may be good proxy measures of overall production activity when measures of production output are not available or have specific shortcomings[6]. The second is in the identification for upstream (vertical) integration, i.e. whether a plant makes an intermediate product or purchases some pre-processed input. This is an important "boundary" issue for the energy footprint of a plant, even when two plants produce identical outputs. The third way is a variation of the second, relating to material "quality." If there are alternative input choices that differ qualitatively and also with respect to energy use then input quality measures can be introduced into the benchmark.

The first way process inputs can be helpful in developing a statistical benchmark of energy use is that inputs like materials, labor, or production hours may be good proxy measures of overall production activity when measures of production output are not available or have specific shortcomings. If a physical measure of output is not readily available and pricing makes the value of shipments a questionable measure of production then physical inputs can be a useful proxy. For some industries the basic material input is so ubiquitous that it makes sense to view energy use per unit of basic input rather than (diverse) outputs. Process inputs may also be useful in measuring utilization, either directly or indirectly. Corn refining is a good example of this approach. The industry uses a ubiquitous inputs, corn. In some industries physical production data is not reported in the Census of Manufacturing but material flows are reported and can be used instead. For example, sand, lime, soda ash, and cullet (scrap glass) are the primary inputs to glass manufacturing.

The second way that process inputs are important for inter-plant benchmarking is when vertical integration is common in a sector but varies in degree from plant to plant. Industries are categorized by the products they produce, but some sectors may face a "make or buy" decision in the way they organize production. A plant may purchase an intermediate product or produce it at the plant as part of a vertically integrated plant. For example, an auto assembly plant may stamp body panels or ship them in from a separate facility. The energy use of these two facilities is not directly comparable. The inter-plant benchmarking approach must account for those "make or buy" decisions in the specific plant configurations. Examples range from food processing, where plant may make juice from concentrate or fresh fruit or paper mills which may purchase market pulp or recycled fibers.

The third way that process inputs are important for inter-plant benchmarking is when differences in material quality may also be related to energy use. For example, if the materials mix to produce a product directly impacts energy uses, then the statistical model can apply different weights to the materials mix in the same manner that it does with product mix. In other words, product/process level differences in energy use can be inferred from the volume and types of materials used in production. To be considered in the statistical normalization, they must be measured on a consistent plant-level basis across the industry. For cement plants the hardness

[6] As discuss in Freeman, et al (1997)

and moisture content of the limestone is hypothesized to influence energy use, but no consistent data is available for this, leaving it the subject of future analysis if data can be collected.

One ubiquitous input is labor. Labor may be helpful in capturing the quasi-fixed nature of energy if there are production slow-downs or non-production periods of operation, but when both labor and energy are being used. In this way labor captures a plant activity level that is related to energy use, even when product output is not being generated. As an empirical issue the statistical significance, or lack thereof, of labor in the analysis can capture this potentially industry specific phenomenon.

Climate

There are many things under the control of a plant or energy manager, but one they cannot control is "the weather." In most manufacturing plants heating, ventilation and cooling (HVAC) contributes to energy demand and weather determines how much is required to maintain comfort. Since the approach used here is annual, seasonal variation does not enter into the analysis, but differences due to the location of a plant and annual variation from the locations norm will play a role. The approach that has been taken for all sectors under study is to include heating and cooling degree days (HDD and CDD) into the analysis to determine how much these location driven differences in "weather" impact energy use**.**

In principle all plants have some part of energy use that is HVAC related, but when the HVAC component of energy use is small relative to total plant consumption the statistical approach may not be able to measure the effect accurately enough to meet tests for reliability. For some sectors weather is a factor in the process, like auto assembly. It a factor because of paint booths and climate control technology need for those systems. Pharmaceuticals manufacturing, where "clean room" production environment is common, is another good example. The climate impact in this sector is only applicable to the "finish and fill" production stage. The more energy intensive chemical preparation stage is not sensitive to climate. Even in industries where the HVAC component is not an obvious or large part of energy use there may be production process related effects that analysis needs to test for. For example, processes that use chillers may be sensitive to CDD (summer) loading. Process heat furnaces may be sensitive to cold outside air so HDD (winter) effects might be included in the model.

Methodology

These case studies do not share a single methodology, but the approaches can be characterized by whether 1) the analysis is linear (in either energy use or energy intensity) or log linear and 2) the assumption about the distribution of the error term.

The underlying approach of these case studies is a regression model of the general form

$$E = f(Y, X; \beta) + \epsilon$$

Where *E* is the measure of energy, *Y* is either production or a vector of production related activities, *X* is a vector of plant characteristics, β is a parameter vector (the normalization factors) and ε is the measure of relative plant efficiency. This is similar to studies of total factor productivity where the LHS is production, the RHS is a production function, and ε is the Solow residual. (Syverson 2011) provides a review of the productivity dispersion literature. (Boyd

2008) provides some of the theoretical connections between the function f(.), the *energy factor requirements function*, and the sub-vector distance function (see (Murillo-Zamorano 2004) for a review of the distance function approaches). The error term is interpreted as a measure of *relative energy efficiency.* The estimates of the distribution of the error term are used to reflect the relative spread of energy performance within a sector.

One variation in the case studies is one imposing the relationship be homogeneous of degree one in a single production variable, y.

$E = y\,(f(X;\beta) + \epsilon)$, or rearranging to an intensity form $E/y = f(X;\beta) + \epsilon$

This approach ignores many issues about the assumption of the distribution of the error term but simply posits the regression model in the intensity form. When the function f(.) is linear in X then this represents the ***linear methodology***. This approach may be appropriate if the activities represented by X are additively independent, e.g. in corn refining the moisture content of one of the primary by-products influences the energy use but has no impact on the energy use of other products.

An alternative approach is the ***log-log methodology,*** where E and Y are in natural logarithms.

$$\ln(E) = (f(\ln(Y), X; \beta) + \epsilon)$$

$$\boldsymbol{ln(E) = a + \sum_{i=1}^{n} b_i\, ln(y_i) + \sum_{i=1}^{m} c_i\, X_i + \varepsilon}$$

The error term is now interpreted as a percent efficiency, rather than absolute levels or intensities as in the linear approach. Whether this assumption of linearity is one of convenience or not, there are only two case studies that use the linear approach; all others use log linear. One advantage of the log linear approach is that it provides an easy estimate of energy returns-to-scale. In other words the sum of the coefficients, *b*, of the activity vector, *Y*, is a measure of whether energy use scales proportionally to total activity. If the sum of the coefficient are close to unity then larger plants do not have an “advantage” in terms of lower energy use. If the sum is less than unity then larger plant use less than a proportional amount of energy. If capacity is included in the Y vector then the case study provides both a short and long run estimate.

The distributional assumptions for the error term can reflect whether the efficiency in the industry is approximately (log) normal of follows a skewed behavior. All the case studies consider the possibility that efficiency is skewed and test the stochastic frontier energy model as in (Boyd 2008). This assumes the error term is composed of two parts

$$\varepsilon = u - v$$

Where

$$v \sim N(0, \sigma_v^2)$$

and energy (in)efficiency, *u,* is distributed according to some one-sided statistical distribution[7], for example gamma, exponential, half normal, and truncated normal. It is then possible to

[7] We also assume that the two types of errors are uncorrelated, i.e. $\sigma_{u,v} = 0$.

estimate the parameters of equation (2), along with the distribution parameters of *u* and *v* using maximum likelihood methods. The approach that is used to estimate these parameters depends on the type of distribution that is used to represent inefficiency. Exponential and truncated normal frontier models can be estimated using relatively conventional maximum likelihood (ML) techniques available in many modern statistical packages. Gamma is a very flexible distribution, but also generates a model that is very difficult to estimate since there is no closed form for the likelihood function and a simulated maximum likelihood has been proposed. A wide range of additional distributional assumptions regarding the heteroscedasticity of either *u* or *v* are also possible. In addition, the treatment of panel data is a significant issue in the application of stochastic frontier, since the inefficiency term is likely to be correlated over time within a plant or firm. (Greene 2002) presents an overview of panel treatments[8]. If empirically there is no evidence of skewness of the error term then the ML estimate of σ_v^2 will be close to zero and the ML estimates are equivalent to OLS.

Since the distribution of the error term, composed or normal, is taken as a measure of efficiency dispersion heteroscedasticity takes on a specific interpretation beyond the usual concerns for parameter standard errors. If the heteroscedasticity is related linearly to some variable in the X vector of the form

$$\sigma_{i,v} = \gamma X_{k,i}$$

or to production, Y

$$\sigma_{i,v} = \gamma Y_i$$

or to the inverse of production

$$\sigma_{i,v} = \gamma \, {}^{1}\!/_{Y_i}$$

then the efficiency distribution and the associated quartiles depends on the size of X or Y. In the last example the distribution of efficiency is larger for low production plants than for high production plants.

Overview of selected case studies

Drawing from the general approach above, Table 1 summarizes the factors that have been included in each of the industry case studies to explain difference in inter-plant energy use. It is clear that each industry is unique in the characteristics that "matter" for energy benchmarking. Twenty of the studies use some type of physical units for activity; of those, 18 have 2 to 5 different sub-product types or use some other information to further characterize product differences, i.e. price or size. Some measure of plant size and utilization is included in 5, but the small number is due more to data limitations, i.e. available plant level capacity information. Person or operation hours are included in 8 industries. In some cases the labor hours *may* be playing a similar role to utilization, i.e. capturing non-production activity that uses energy. About half of the sectors include process inputs, either as a ubiquitous measure of input, e.g. corn in corn refining or scrap in minimills, or in the form of raw vs. preprocessed inputs, e.g. fresh fruit vs concentrate in Juice production or virgin vs. recycled fiber in paper production. The

[8] No case studies these panel treatments here since they only have, at most, a few years of data for any given plant in an unbalanced panel.

selection of inputs is based in part on data availability, but then only included when the estimated effect is of reasonable size and statistically significant.

Table 2 further describes the statistical form of the models. Seven sectors exhibit a skewed distribution of energy intensity and are modeled as stochastic frontiers; the rest are best approximated as log normal, i.e. the percentage difference from average performance are "bell shaped." Two of the log normal models exhibit heteroscedastic distributions with variance declining wrt production volumes. The earliest year of data for a study is 2002. This is largely driven by the data available when the analysis was conducted. 2007 is the most recently available data from Economic Census[9]. Sectors that use industry or trade association provided data tend to have more recent benchmark years. For the less energy intensive industries using Census of Manufacturing (CM) data, the energy content of the fuels is imputed using cost data and state level energy prices. This is done since the sample sizes in the Manufacturing Energy Consumption Survey (MECS) are too small to meet disclosure requirements. For industries with larger MECS samples the more detailed energy information is used directly. Sample sized vary depending on the industry, although these sample sizes should be viewed as a "fairly complete" count of all the plants in that sub-sector. Although some data is dropped due to missing variables from incomplete reporting or other data quality screens such as for extreme outliers.

[9] As of this draft the 2012 EC was not yet available in the RDC network

Table 1: ES-EPI Benchmarks Inputs, by Industry and Sub-Sector

Focus industries	Product mix	Units	Inputs	Size or capacity	External	Other
Cement (V 2.0)	3 product types	Tons	-	Capacity & # of Kilns	Utilization	Person hours
Corn Refining (V 2.0)	5 product types	Bushels	Corn	Capacity	Utilization	Feed moisture
Dairy - Fluid Milk *	6 product types	Gallons	Whole milk	-	CDD	Person hours
Dairy - Ice cream *	4 product types	Gallons	2 types	-	CDD	Person hours
Ethyl Alcohol **	Single	Gallons	-	-	-	-
Food - Juice (Canned)	4 x 2 product types	Gallons	2 types	-	-	-
Food - Frozen Fried Potatoes	Single	pounds	-	-	-	Warehouse (frozen)
Food - Tomato products **	1 sub-product type		2 types	-	-	Person hours
Baking - Cookies & Crackers	3 product types	pounds	-	-	-	-
Baking - Bread & rolls *	5 Product types	tons	Raw dough	-	Weather TBD	Freezers
Glass – Flat	-	pounds	Sand	-	-	-
Glass – Container	Price	pounds	Sand, Cullet	-	-	-
Iron and Steel - Integrated *	3 stages of production	tons	-	Furnace capacity	Utilization	-
Iron and Steel – Minimills *	Price	tons	Scrap	Furnace capacity	Utilization	-
Metal casting - Iron *	4 product types	Tons, price	-	-	HDD	Person hours
Metal casting - Investment steel *	-	hours	-	-	-	Person hours
Metal casting - “Other” steel *	3 product types	$	-	-	-	-
Metal casting – aluminum *	3 product types	Tons, price	-	-	HDD	-
Motor Vehicle - Assembly V2.0	vehicle size	of vehicles	-	-	Weather, Utilization	Air Tempering
Pharmaceuticals	3 activity types *	%	-	Facility size (ft2)	Weather, Utilization	Operation hours
Printing - Lithograph *	6 product types	$	2x3 types	-	Weather TBD	-
Pulp Mills	3 product types	tons	2 types	-	-	Water treatment
Paper & Board Integrated Mills	3 product types	tons	3 types	-	-	Water treatment, bleaching chemicals

Focus industries	Product mix	Units	Inputs	Size or capacity	External	Other
Ready Mix Concrete *	**2 activities**	**Tons,miles**	**-**	**-**	**-**	**-**

* Under Industry Review, ** Preliminary

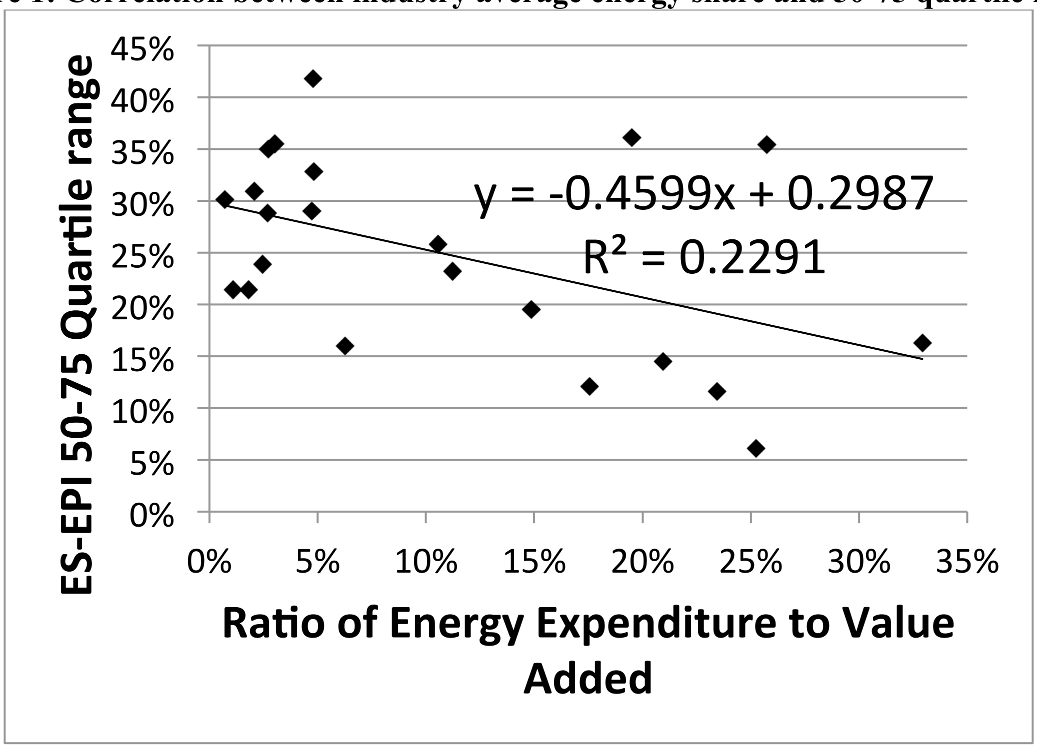
Table 2: ES-EPI Benchmarks Model Details, by Industry and Sub-Sector

Focus industries	Model	Year	# of plants	Data source	RTS	75 to 50th
Cement (V 2.0)	log normal (heteroscedatic)	2000-2008	96	1.0	Variable SR 0.92 LR	-6.1%
Corn Refining (V 2.0)	half normal frontier	2004-2009	37	Industry	Variable SR Constant LR	-14.5%
Dairy - Fluid Milk *	log normal	2002	258	CM	0.85	-29.0%
Dairy - Ice cream *	log normal	2002	89	CM	1.05	-23.9%
Ethyl Alcohol **	log normal	2007	111	CM	0.70	-35.4
Food - Juice (Canned)	log normal	2002	44	CM	0.84	-41.8%
Food - Frozen Fried Potatoes	log normal	2002	27	CM	0.91	-16.0%
Food - Tomato products **	log normal	2002	40	CM	1.11	-43.7%
Baking - Cookies & Crackers	log normal	2002	64	CM	0.71	-30.9%
Baking - Bread & rolls *	log normal	2007	207	CM	0.91	-28.8%
Glass – Flat	log half normal frontier	2002	38	CM, MECS	Variable	-16.3%
Glass – Container	log normal	2002	62	MECS	1.03	-11.6%
Iron and Steel - Integrated *	log exponential frontier	2005-2009	12	Industry	0.72 SR 0.99 LR	9.2%
Iron and Steel – Minimills *	log normal	2002	39	CM, MECS		-12.1%
Metal casting - Iron *	log normal	2006	83	CM, MECS	1.06	-23.2%
Metal casting - Investment steel *	log half normal frontier	2007	51	CM	1.03	-32.8%
Metal casting - "Other" steel *	log normal	2007	59	CM	Variable	-25.8%
Metal casting – aluminum *	Log normal (heteroscedatic)	2007	290	CM, MECS	1.05	-28.3%
Motor Vehicle - Assembly V2.0	Gamma frontier	2003-2005	33	Industry	Variable SR Constant LR	-21.4%
Pharmaceuticals	log half normal frontier	2004-2006	61	Industry	Variable SR 0.98 LR	-30.1%
Printing - Lithograph *	Log half normal	2007	775	CM	1.0	-35.0%
Pulp Mills	log normal	2002	28	CM, MECS	1.05	-36.1%
Paper & Board Integrated Mills	log normal	2002	99	CM, MECS	0.71	-19.5%
Ready Mix Concrete *	log normal	2008-2009	62	NRMCA	0.83 SR 0.89 LR	-35.5%

* Under Industry Review, ** Preliminary

The last column labeled 75 to 50th represents the third quartile range, i.e. percent difference of the 75th percentile, i.e. the ENERGY STAR certified plant level, and the average or median performance, the 50th percentile. This ranges from as low as 6% to nearly 44%. Figure 1 compares this third quartile range to the industry average share of energy cost to value added. This cost share reflects how "important" energy is in the sector. We see a clear correlation between high cost shares and the range of performance. This makes sense since industries with higher relative energy costs would put more effort into management of those costs. There are outliers in this relationship, however. They include pulp mills and ethanol (dry mill) plants. The latter is a preliminary estimate. The result for pulp mills may suggest the need for additional scrutiny. However, the EPI uses net purchased energy and pulp mills provide a large amount of internally generated power from black liquor and CHP. There may actually be a wide range of practices in terms of net purchased energy in this sector than for other energy intensive ones.

Figure 1: Correlation between industry average energy share and 50-75 quartile range



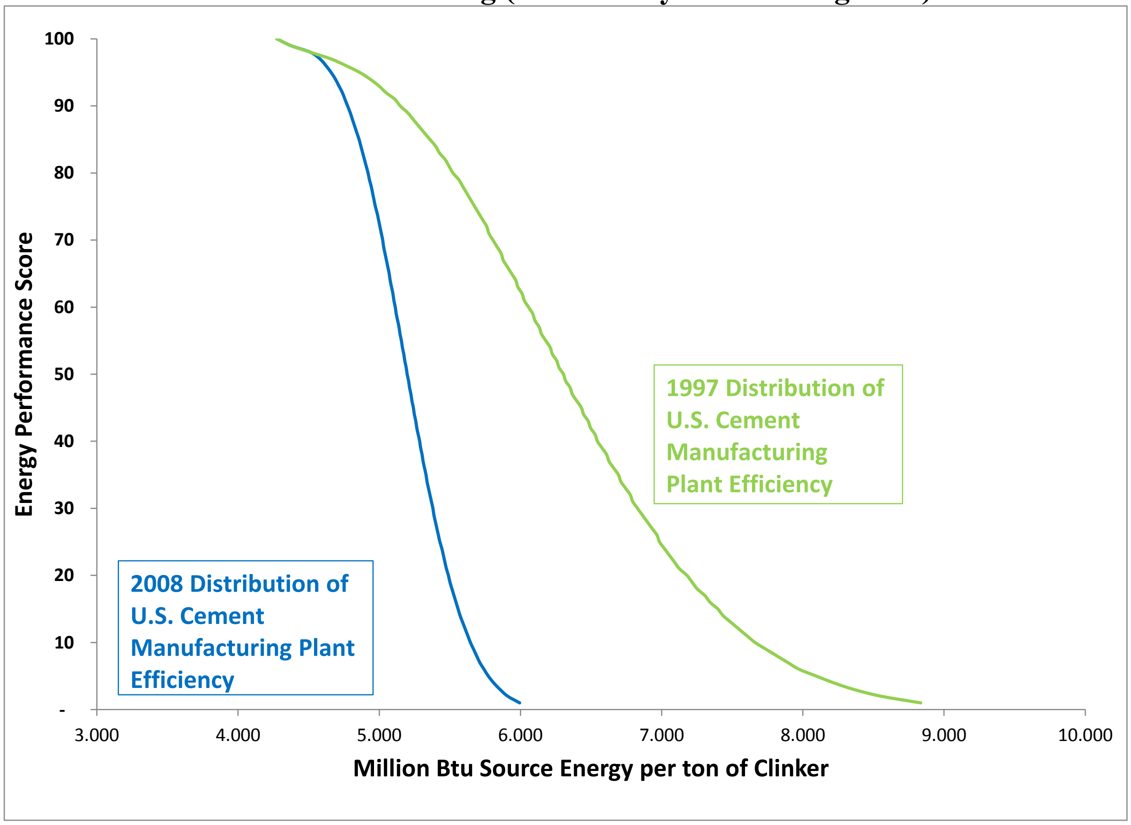
The second to the last column is the estimate of returns to scale wrt energy use. Constant returns is the most common; only 8 studies suggest some economies of scale. The studies that use linear forms impose constant returns in the long run and allow for variable returns in the short run. Some studies use a variable returns specification by including second order terms in the activity variables or a measure of capacity utilization, i.e. production/capacity, in the analysis.

Updates for benchmark year for three ES-EPI

In the 2010, first three completed case studies; Auto assembly, Cement, and Corn refining (see (Boyd 2005, Boyd 2006, Boyd 2008) for detailed descriptions of the earlier models) began to be updated. Comparing the old benchmark with the new benchmark reveals information about how these three, very different industries have changed over time. Since the analysis reveals both the general level and range of energy performance the comparison focuses on how much the change in the "best practice" and the change in the range of performance contribute to the overall reduction in energy use in the sector (see (Boyd and Zhang 2012), (Boyd and Delgado 2012), and (Boyd 2010) for the details of the updates).

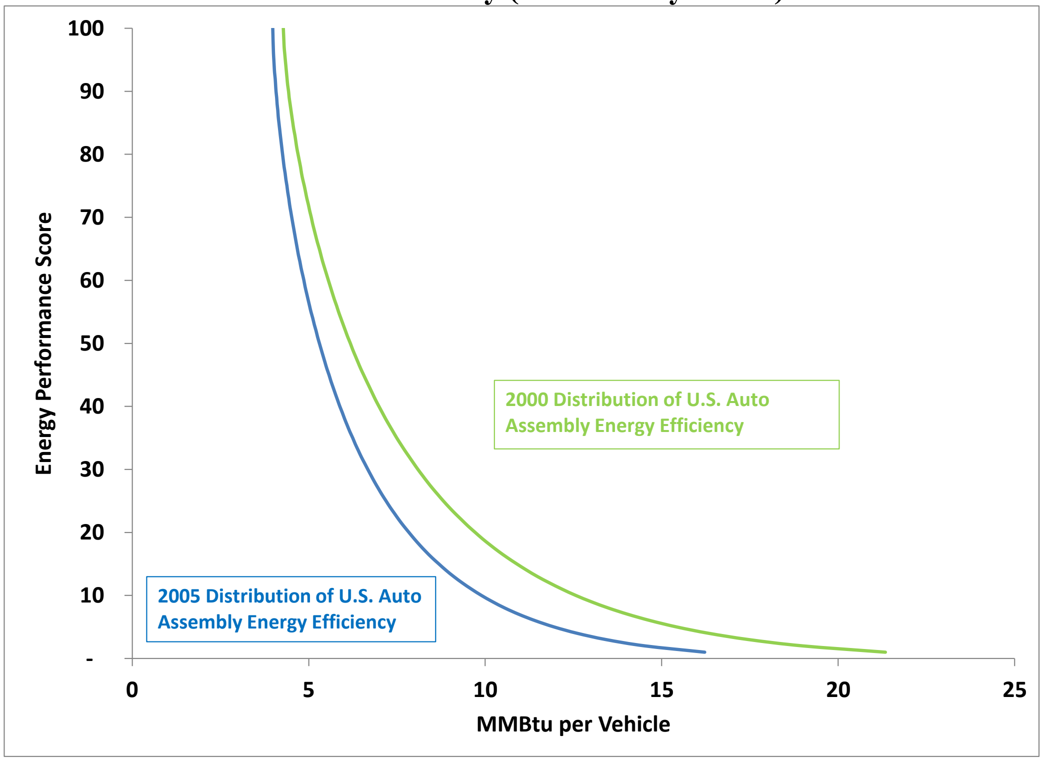
For the cement industry, if one computes the ratio of total energy costs to total value of shipments (adjusted for inflation) in 1997 and 2007 from data collected in the Economic Census, one would conclude that this measure of energy intensity has fallen ~16%, from 0.184 to 0.158. Aggregate data may also give the impression that all plants have made the same steady improvements. The picture that emerges from our plant level statistical analysis is somewhat different and more subtle (figure 2); poorer-performing plants from the late 1990s have made efficiency gains, reducing the gap between themselves and the top performers, whom have changed only slightly. The results from this study focus on energy efficiency and controls for other structural changes in the industry, e.g., increases in average plant size, which also tend to lower energy use. Our estimate of the overall energy efficiency improvement in the 96 plants in our database represents a 13% percent change in total source energy and the source of these changes is clearly not uniform.

Figure 2: Comparison of Two Benchmark Distributions of Energy Efficiency in Cement Manufacturing (source: Boyd and Zhang 2012)



Results for the auto assembly industry are similar, but less dramatic (figure 3). There are two sources of improvement, the changes in the industry energy frontier, i.e. "Best Practices" and technology, and the changes in efficiency, i.e. whether plants are catching up or falling behind. The results suggest that slightly more than half of the improvement is changes in efficiency, which have slightly outpaced changes in the frontier. The combined effect when evaluated against the over 7 million vehicles produced in 2005 by the plants in our study implies in a reduction of 11.6%, or 1462 million lbs of CO_2, attributable to changes in observed industry energy efficiency practices.

Figure 3: Comparison of Two Benchmark Distributions of Energy Efficiency in Auto Assembly (Source: Boyd 2010)



The change in the distribution of energy efficiency for a representative corn refining plant is shown in figure 4. If we multiply this plant-specific change in energy intensity by the level of corn input production for each plant operating in the industry in 2009, and total across all plants, we compute a reduction of 6.7 trillion Btu in annual energy use. Relative to an average annual total source energy consumption of 155 trillion Btu in 2009 for all the plants in our data set, this represents about a 4.3% reduction in overall energy use by this industry. When energy-related greenhouse gas emissions are considered, this represents an annual reduction of 470 million kg of energy-related CO_2 equivalent emissions from improved energy efficiency. The change in performance from these three industry are all quite different. Cement reflects the case where best practice has changed very little, but "catching up" comprises the main source of improvements. Corn refining is at the opposite end of this spectrum, where there are substantial changes in the best plants, but laggards remain or in some sense are even falling behind by failing to keep pace. The auto assembly plants are a mixture of changes in best practice and some modest "catching up". These benchmark updates also reflect different time periods. When we compute the average annual change from the total reduction in energy use for each sector we see that the auto industry has made the greatest strides (see table 3).

Figure 4: Comparison of Two Benchmark Distributions of Energy Efficiency

in Wet Corn Refining (source: Boyd and Delgado 2012)

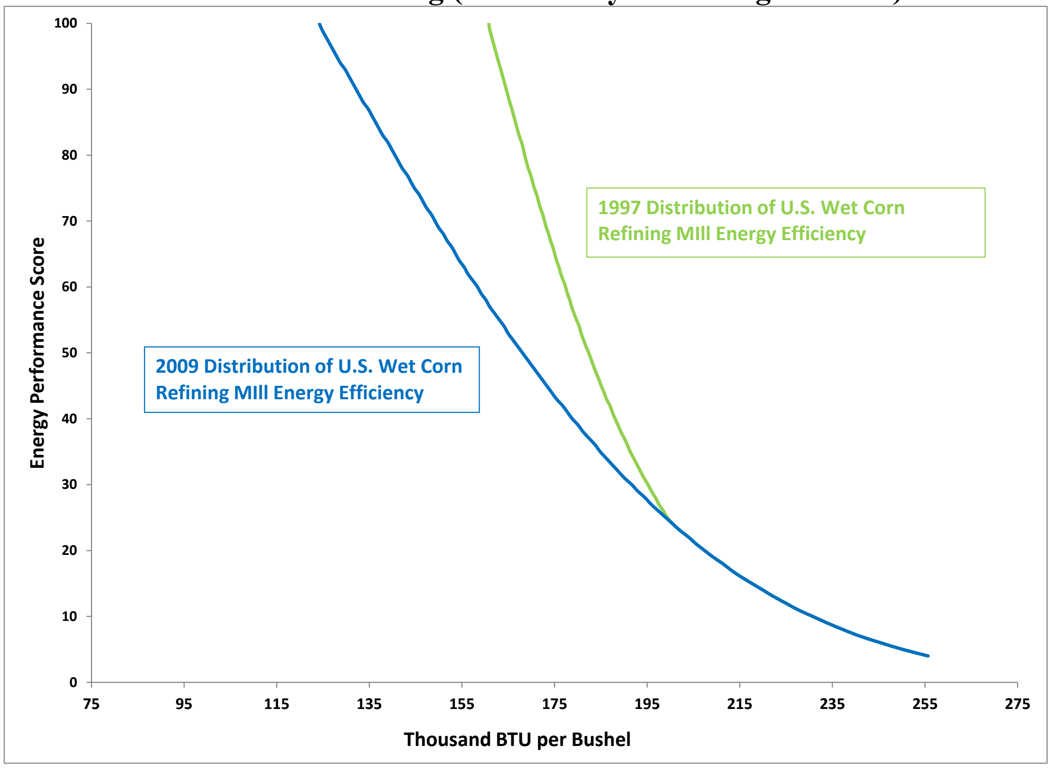


Table 3: ES-EPI Benchmarks Updates: Rate of change by Industry

Sector	New benchmark Year	Old Benchmark Year	Time period	Total reduction	Average annual change
Auto	2005	2000	5	12.0%	2.3%
Cement	2008	1997	11	13.0%	1.2%
Corn	2009	1997	12	4.3%	0.4%

Conclusions

The objective of these case studies is for developing sector-specific energy performance benchmarks and to create a tool that would motivate companies to take actions to improve the energy efficiency of their plants and ultimately help reduce greenhouse gas emissions in the industrial sectors benchmarked. As of December 2012, EPA had published 11 EPIs, awarded 120 ENERGY STAR plant labels, and engaged an additional twelve industrial sectors and subsectors in the EPIs development process (see Table 3). Compared to average plants (EPS score of 50) EPA estimated in 2011 that plants earning the ENERGY STAR have saved an estimated 314,190,357 MMBtus. [10]Companies using the ES-EPI report that they find the tools valuable and beneficial for evaluating current performance and setting efficiency goals. Many companies report they have incorporated the ES-EPI into their energy management programs and have made achieving ENERGY STAR certification as an objective.

[10] US EPA (2012)

Initially, there was industry skepticism that a whole-plant benchmark could be developed using statistical case studies. Skeptics largely believed that each plant is too "unique" for whole plant comparisons to be made. However, both the process and method used to develop the case studies has helped change skeptics participating in the industrial focus process into supporters. The *process* of engaging the industry in the development of the case studies has been critical. By developing the case studies in a transparent, objective, and collaborative process, industry participants were directly involved in the design and review from the beginning. This process enabled the identification of potential factors for inclusion in the regression analysis, receive timely feedback on draft results, quickly address concerns, and ultimately ensure a high degree of support and "buy-in" for the tool. By using a benchmarking method based on actual operational data and that allowed for controls to address industry specific differences between plants, concerns were overcome that industrial plants are too heterogeneous, even within a specific sub-sector, to be able to benchmark.

The availability of sector-wide energy and production data through the US Census Bureau was critical for the analysis. One the greatest barrier to any benchmarking exercise is inadequate or unrepresentative data. The case studies has benefited from the robust industrial energy and production data collected by the US Census through the Census of Manufacturing (CM) and the Manufacturing Energy Consumption Survey (MECS). The availability of this data for use in developing the statistical models has been critical to ensuring the early success of the ENERGY STAR industrial benchmarking program. First, it provided EPA with the ability to develop the benchmarks without having to undertake a data collection. Second, by working with Census data, which has strict confidentially requirements, the ENERGY STAR team was able to build trust amongst industry participants that the company specific data used for benchmarking would be kept confidential and would not be shared with either focus participants and the EPA. While some of the more recent case studies have drawn on data provided by the industry, the availability and quality of the CM and MECS data enabled ENERGY STAR to successfully develop the first case studies and demonstrate that whole-plant energy performance benchmarking is possible.

The process of developing EPIs has uncovered new insights into energy use and the drivers of efficiency within the sectors benchmarked. Additionally, the establishment of industry baselines has enabled EPA to visualize the range of performance within a sector. Visualizing the distribution of performance offers important information for policy makers and others interested in promoting efficiency or reducing GHG emissions from specific industrial sectors. The slope of the baseline curve generated by the EPI can help policy makers and others evaluate what action is needed to improve the performance of the industry. For example, sectors with steep baseline curves and distributions indicate that the opportunities for improving energy efficiency through existing measures may be limited. These sectors should be considered for R&D investments to develop new technology that can create a step change in the level of performance. Additionally, these sectors may face greater difficulties reducing their GHG emissions through existing energy management measures. Whereas sectors with flatter curves indicate that more opportunities are available through existing technologies and practices. In these sectors, there is a greater distribution of performance, which usually suggests that existing energy management measures and investments can improve performance.

The process of benchmarking and re-benchmarking a sector provides further insights into the improvement potential of the industry over time. Understanding how the distribution of energy performance in a sector is changing or not changing can provide valuable information for policy makers as well as business leader in developing strategies to drive future performance gains.

The approach and method used by ENERGY STAR to benchmark whole-plant energy performance has potential applicability to other sustainability metrics, such as water and waste, as well as sub-systems within plants. While developing such benchmarks is beyond the scope of the ENERGY STAR program, several companies participating in the Industrial Focus process have recently initiated an independent effort that applies the ENERGY STAR benchmarking approach to process lines within the plant and to non-energy measures such as water. If successful, the results of this effort will break new ground in advancing the field of energy performance and sustainability benchmarking.

Table 3: Status of Case Studies (published or under review)

Focus industries	Status
Cement (V 2.0)	**(Boyd and Zhang 2012)**
Corn Refining (V 2.0)	(Boyd and Delgado 2012)
Dairy - Fluid Milk	Under industry review
Dairy - Ice cream	Under industry review
Ethyl Alcohol	Under industry review
Food - Juice (Canned)	(Boyd 2011)
Food - Frozen Fried Potatoes	(Boyd 2011)
Food - Tomato products	(Boyd 2011)
Baking - Cookies & Crackers	(Boyd 2011)
Baking - Bread & rolls	Under industry review
Glass – Flat	(Boyd 2009)
Glass - Container	(Boyd 2009)
Iron and Steel - Integrated	Under industry review
Iron and Steel – Minimills	Under industry review
Metal casting - Iron	Under industry review
Metal casting - Investment steel	Under industry review
Metal casting - “Other” steel	Under industry review
Metal casting - Aluminum	Under industry review
Motor Vehicle - Assembly (V2.0)	(Boyd 2014)
Pharmaceuticals	(Boyd 2009)
Printing - Lithograph	Under industry review
Pulp Mills	(Boyd and Guo 2012)
Paper & Board Integrated Mills	(Boyd and Guo 2014)
Ready Mix Concrete	Under industry review

References

Boyd, G. (2009). Development of a Performance-based Industrial Energy Efficiency Indicator for Glass Manufacturing Plants. Washington DC, US EPA**:** 26.

Boyd, G. (2011). Development of Performance-based Industrial Energy Efficiency Indicators for Food Processing Plants Washington DC, US EPA.

Boyd, G. (2012). A Statistical Approach to Plant-Level Energy Benchmarks and Baselines: The Energy Star Manufacturing-Plant Energy Performance Indicator. Carbon Management Technology Conference. Orlando, Florida USA.

Boyd, G., E. Dutrow and W. Tunnessen (2008). "The evolution of the ENERGY STAR® energy performance indicator for benchmarking industrial plant manufacturing energy use." Journal of Cleaner Production **16**(6): 709-715.

Boyd, G. and Y. Guo (2012). Development of Energy Star® Energy Performance Indicators for Pulp, Paper, and Paperboard Mills. Washington DC, US EPA**:** 30.

Boyd, G. and G. Zhang (2012). "Measuring improvement in energy efficiency of the US cement industry with the ENERGY STAR Energy Performance Indicator." Energy Efficiency: 1-12.

Boyd, G. A. (2005). Development of a Performance-based Industrial Energy Efficiency Indicator for Automobile Assembly Plants. Argonne IL, Argonne National Laboratory**:** May 2005.

Boyd, G. A. (2005). "A Method for Measuring the Efficiency Gap between Average and Best Practice Energy Use: The ENERGY STAR Industrial Energy Performance Indicator." Journal of Industrial Ecology **9**(3): 51-65.

Boyd, G. A. (2006). Development of a Performance-based Industrial Energy Efficiency Indicator for Cement Manufacturing Plants. Argonne IL, Argonne National Laboratory.

Boyd, G. A. (2008). "Estimating Plant Level Manufacturing Energy Efficiency with Stochastic Frontier Regression." The Energy Journal **29**(2): 23-44.

Boyd, G. A. (2009). Development of a Performance-based Industrial Energy Efficiency Indicator for Pharmaceutical Manufacturing Plants. Druham NC, Duke University. report to the Environmental Protection Agency

Boyd, G. A. (2010). Assessing Improvement in the Energy Efficiency of U.S. Auto Assembly Plants. Duke Environmental Economics Working Paper Series, Nicholas Institute for Environmental Policy Solutions.

Boyd, G. A. (2014). "Estimating the changes in the distribution of energy efficiency in the U.S. automobile assembly industry." Energy Economics **42**(0): 81-87.

Boyd, G. A. and C. Delgado (2012). Measuring Improvement in the Energy Performance of the U.S. Corn Refining Industry. duke environmental economics working paper series. Durham NC, Nicholas Institute For Environmental Policy Solutions**:** 1-18.

Boyd, G. A. and Y. Guo (2014). "An Energy Performance Indicator for integrated paper and paperboard mills: A new statistical model helps mills set energy efficiency targets." Paper360 March/April: 26-28.

Boyd, G. A. and W. Tunnessen (2007). Motivating Industrial Energy Efficiency through Performance-based Indicators. ACEEE Summer Study on Energy Efficiency in Industry Improving Industrial Competitiveness: Adapting To Volatile Energy Markets, Globalization, And Environmental Constraints White Plains, New York.

Farrell, M. J. (1957). "The Measurement of Productive Efficiency." Journal of the Royal Statistical Society. Series A (General) **120**(3): 253-290.

Freeman, S. L., M. J. Niefer and J. M. Roop (1997). "Measuring industrial energy intensity: practical issues and problems." Energy Policy **25**(7-9): 703-714.

Greene, W. H. (2002). Fixed and Rondom Effects in Stochastic Frontier Models. New York Univesity Economics department working paper.

Murillo-Zamorano, L. (2004). "Economic efficiency and frontier techniques." Journal of Economic Surveys **18**(1): 33-77.

Syverson, C. (2011). "What Determines Productivity?" Journal of Economic Literature **49**(2): 326-365.